

A spatial high-resolution model of the dynamics of agricultural land use

Luis Ambrosio Flores , Luis Iglesias Martínez , Carmen Marín Ferrer , Valero Pascual Gallego ,
Arturo Serrano Bermejo

Departamento de Economía y Ciencias Sociales Agrarias, Escuela Técnica Superior de Ingenieros Agrónomos, Universidad Politécnica de Madrid, Avda Complutense sn 28040, Madrid, Spain

Departamento de Explotación de Recursos Minerales y Obras Subterráneas, Universidad Politécnica de Madrid, Spain

Departamento Interdisciplinar de Desarrollo Sostenible, Universidad Católica de Ávila, Spain

Abstract

An area model is presented for agricultural land use, based on a generalized linear mixed model. This model is spatially explicit and dynamic and, although it uses aggregated data, allows for heterogeneity of behavior among individual farmers. The parameters of the fixed component of the model are obtained using an estimation equations approach, and the structure of spatiotemporal correlation is assessed using empirical semivariograms. The model is illustrated using as an example the dynamics of agricultural land use in the Lower Guadalquivir area in Spain. A simulation study indicates that the model gives poor results if the heterogeneity of individual behavior and spatiotemporal autocorrelation are ignored.

JEL classifications: C35, C53, C61, Q15, R14

Keywords: Aggregation ecological fallacy; Generalized linear mixed models; Semivariogram models for spatiotemporal autocorrelation structures; Spatial and dynamic choice models

1. Introduction

Many of the greatest environmental problems (such as desertification or loss of biodiversity) are linked closely to land use. Thus, a better knowledge of how land use changes can be useful for the formulation and implementation of environmentally sensitive policies and practices. Since agriculture constitutes one of the largest proportions of land use worldwide, this article focuses on modeling the dynamics of agricultural land use. Models of changes in agricultural land use, together with environmental indicators related to water consumption, soil erosion, the use of herbicides and pesticides, and greenhouse gas emissions, can contribute to the introduction of better policies for promoting sustainable agriculture.

Advances in territorial observation technology (using satellite imagery) and in geographical information systems (GIS) have made area models a powerful and useful tool for investigating changes in land use. Recently, numerous models have been proposed that use data from small territorial units that are collected using a sampling scheme (Miller and Plantinga, 1999;

Plantinga, 1996; Wu and Segerson, 1995; Wu et al., 2004), or by using satellite images (Müller and Zuller, 2002; Munroe et al., 2002). In these cases, changes in land use are explained in terms of the relative economic productivity of the different uses.

As a result of integrating biophysical and socio-economic data at varying spatial resolutions, it is necessary in these models to create territorial microunits that facilitate comparison among the data sets used (Gotway and Young, 2002). In order to illustrate the model developed in this article, a data set will be used that draws on a sampling scheme proposed by the Joint Research Centre of the European Union (EU) (Gallego et al., 1994). The territorial microunit, A_i , will be a 10 km by 10 km block. All data will be aggregated to that level. Two factors will be considered as having the most influence on farmer behavior in this model: (i) grant aid from the EU for agricultural production and (ii) the physical condition of the territorial microunit. EU grant aid is officially defined at the municipal level and aggregated to the territorial microunit scale using a weighted average. The weighting coefficient is derived from the area of the municipality within the block. The physical condition of the microunit is evaluated using the smallest homogeneous soil unit registered on soil maps and is aggregated to the level of the territorial microunit as a weighted average. The weighting

coefficient is derived from the area of the homogeneous unit within the block.

The values that result from each microunit are represented by an aggregate value from a number (or conglomerate) of different farmers (i.e., all the farmers operating within a microunit). That being so, a fundamental issue arises, namely, whether the relationships between land use and associated explanatory variables estimated using the aggregated data at microunit level are consistent with the actual relationships at the level of the individual farmers. The general conditions for this consistency are established in aggregation theory (Ijiri, 1971). In studies where this issue has arisen explicitly, it is known as the “aggregation ecological fallacy” (Gotway and Young, 2002). Unfortunately, in many applications, it is difficult to quantify such general conditions. Another approach to the issue of consistency is to model both, the individual and the aggregated data, in the same way. This is the approach followed in this article.

Most suggested area models are homogeneous in the sense that they assume the existence of a “model farmer” and ignore the heterogeneity of behavior among individual farmers. An economic analysis of the limitations of homogenous models can be found in Stoker (1993), together with a review of the approaches that take the heterogeneity between individuals of the same conglomerate into consideration at the time that aggregated data are modeled. This article proposes a new method for modeling aggregated data that takes into account the heterogeneity of behavior among farmers.

Patterns of land use emerge from interactions between human activity and the environment. The context in which such interactions take place lies at the boundary between the natural and social sciences. In this complex context, many factors cannot be observed (Veen and Otter, 2001). However, in the human behavior model, context is a factor of tremendous importance (Luke, 2004). With this in mind, an approach is suggested herein that accounts for the effect of unobservable contextual factors. The purpose is to explain farmers’ decisions about land use in terms of both observable indicators of behavior (explanatory variables) and unobservable contextual factors. The effect of the interactions of the individual farmers with these latter factors is modeled by introducing a latent variable (Skrondal and Rabe-Hesketh, 2004). Attention is focused on modeling the variability and spatiotemporal autocorrelation of this latent variable. This is a critical factor in dynamic models for distinguishing true intertemporal dependence between the different states of the modeled process from spurious dependence caused by temporal autocorrelation (Hsiao, 2003).

A common approach to modeling the dynamics of land use at high spatial resolution uses Markov chains within the framework of cellular automata (Baltzer et al., 1998; Irwin and Geoghegan, 2001; Luijten, 2003). Alternatively, a maximum entropy scheme can be used (Howit and Reynaud, 2003). However, this approach is complicated to implement. In this article, the dynamics of land use are modeled using a spatially explicit generalized linear model (GLM) in a generalized estimating equations (GEE) framework. This is simpler than using non-stationary Markov chain analysis. This approach unifies those

used in previous studies, because a Markov chain of any order can be specified as a GLM model (Fahrmeir and Kaufmann, 1987) and for well-defined problems maximum entropy is reduced to a logit model of the GLM family (Golan et al., 1996, p. 32; Miller and Plantinga, 1999).

The predictor of land use will be defined on the basis of the suggested model, along with an estimator of the error of the predictions. Together with the GIS tools, this predictor allows maps for the prediction of land use to be generated with the required spatial resolution. These maps may be useful at a local level for assessing the effect of different policies on the spatial distribution of land use and on natural resources.

2. The model

Usually, area models select one type of land use from various alternatives. In some previous studies (Alig, 1986; Plantinga et al., 1990; Stavins and Jaffe, 1990; White and Fleming, 1980) only two categories of use were considered $K = 2$. In those studies, a binomial logit homogeneous model was used to establish a relationship between (i) the probability, p_{kit} , that the land in the unit of territory A_i is assigned to the category of use ($k = 1, 2$) and (ii) the associated explanatory variables (Lichtenberg, 1989; Parks and Murray, 1994). In this article, a heterogeneous multinomial model ($K > 2$) is considered that admits heterogeneity of behavior between individual farmers. The basis for this is an individual model of behavior for farmers. An area model is defined that aggregates this individual model of behavior to the level of territorial microunits.

The same territory B is considered to be partitioned into M farms and N microunits $\{A_i; i = 1, 2, \dots, N\}$ so that $\bigcup_{i=1}^N A_i = B$ where i is a vector of coordinates that locates A_i within B with respect to a given reference system. This may be denoted as $\{I_{ij}; i = 1, 2, \dots, N; j = 1, 2, \dots, M\}$, which represents all territory units that result from the intersection of both partitions, where I_{ij} is the surface of the j th farm in the i th microunit. Hence, $\bigcup_{j=1}^{M_i} I_{ij} = A_i$ where M_i is the number of farmers operating in the microunit of territory A_i . Fig. 1 illustrates the difference between the farms and the microunits.

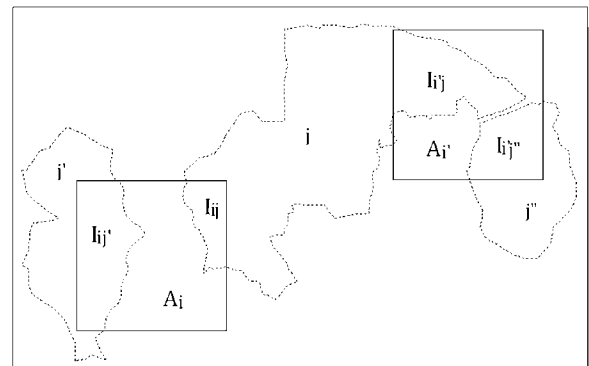


Fig. 1. Partition of the territory in farms and microunits.

The surface areas occupied by each category of use within the microunit of territory A_i in the t th time period are found in the vector, $\underline{Z}_t(A_i) = [Z_{1t}(A_i) Z_{2t}(A_i) \cdots Z_{kt}(A_i) \cdots Z_{Kt}(A_i)]^T$, where $Z_{kt}(A_i)$ is the area occupied by the k th category of use. The surface areas occupied by each category of use to the level of the j th farm are $\underline{Z}_t(F_j) = [Z_{1t}(F_j) Z_{2t}(F_j) \cdots Z_{kt}(F_j) \cdots Z_{Kt}(F_j)]^T$, where $Z_{kt}(F_j)$ is the surface of the k th category of use in the j th farm in the t th period. At the regional macrolevel, the vector of areas for each category of use is $\underline{Z}_t = [Z_{1t} Z_{2t} \cdots Z_{kt} \cdots Z_{Kt}]^T$, where $Z_{kt} = \sum_{i=1}^N Z_{kt}(A_i) = \sum_{j=1}^M Z_{kt}(F_j)$.

We specify an area model at the micro $\underline{Z}_t(A_i)$ and macro levels \underline{Z}_t , using data at the level of the microunits of territory, i.e., data on $\underline{Z}_t(A_i)$ (for $i = 1, 2, \dots, N$ and $t = 1, 2, \dots, T$) and on explanatory variables. The data on land use at the level of the microunits is aggregated in the form $Z_{kt}(A_i) = \sum_{j=1}^{M_i} Z_{kt}(I_{ij})$, where $Z_{kt}(I_{ij})$ is the surface assigned to the k th category of use by the manager of the j th farm in the period t . The essential characteristic of our approach is to model the data at the level of the microunits, $Z_{kt}(A_i)$, using an individual model on the choice of $Z_{kt}(I_{ij})$ on behalf of the manager of the j th farm.

2.1. The subject-specific model

The choice of the individual j th farmer, $Z_{kt}(I_{ij})$, is considered as a realization of a random variable. Its expected value, $E Z_{kt}(I_{ij})$, is specified by means of the GLM $E Z_{kt}(I_{ij}) = g^{-1}(\underline{x}_{kt}^T(I_{ij}) \underline{\beta}_{kijt})$, where $g(\cdot)$ is the link function, $\underline{x}_{kt}(I_{ij})$ is a $(q \times 1)$ vector of explanatory variables, and $\underline{\beta}_{kijt}$ is a $(q \times 1)$ vector of parameters that represents the response of the j th farmer to changes in the explanatory variables.

To allow for individual behavior (heterogeneity among individual farmers as well as interactions of individual farmers with contextual factors), the individual farmer's response is specified as $\underline{\beta}_{kijt} = \underline{\beta}_k + \underline{\delta}_{kijt}$, where $\underline{\beta}_k$ is assumed to be fixed and $\underline{\delta}_{kijt}$ is assumed to be random. Zeger et al. (1988) coined the term "subject-specific effects" for $\underline{\beta}_k$, which represents the response of an average farmer (i.e., a farmer for whom the realization of the individual parameter vector, $\underline{\delta}_{kijt}$, takes its expected value 0) to changes in the explanatory variables. $\underline{\delta}_{kijt} = \underline{\beta}_{kijt} - \underline{\beta}_k$ represents the concrete deviation of the response of individual farmer j compared to the average farmer's response, in location i and time t (the context).

The important thing here is the model for the conditional expectation of $Z_{kt}(I_{ij})$, given individual behavior. The random effect due to (unobserved) individual behavior, $\underline{\delta}_{kijt}$, will be introduced through a latent variable (Robinson, 1991; Stoker, 1993), $\underline{\delta}_{kijt} = \tilde{\underline{x}}_{kt}^T(I_{ij}) \underline{\delta}_{kit}$, where $\underline{\delta}_{kit}$ is a $(q^* \times 1)$ random vector with zero mean and \underline{D} covariance matrix and $\tilde{\underline{x}}_{kt}(I_{ij})$ is a $(q^* \times 1)$ known vector of known random effect variables, some or all of which could be equal to the explanatory variables in $\underline{x}_{kt}(I_{ij})$. The model for this conditional expectation, $\mu_{kijt}(\delta_{kijt}) = E[Z_{kt}(I_{ij}) | \underline{x}_{kt}(I_{ij}), \delta_{kijt}]$, is the log-linear mixed model given by $\mu_{kijt}(\delta_{kijt}) = \exp(\underline{x}_{kt}^T(I_{ij}) \underline{\beta}_k + \delta_{kijt})$, where $\underline{x}_{kt}^T(I_{ij}) \underline{\beta}_k$ is the fixed effect of the explanatory variables (be-

havioral indicators) on the expected response, $\mu_{kijt}(\delta_{kijt})$, given a value of the latent variable, δ_{kijt} . Note that $\mu_{kijt}(\delta_{kijt})$ is conditioned to given behavioral indicator values, $\underline{x}_{kt}(I_{ij})$, although we have suppressed the conditioning for simplicity in this case.

2.2. The population-average model

Instead of the conditional mean $\mu_{kijt}(\delta_{kijt})$, the marginal mean $\mu_{kijt} = E_{\delta_{kijt}} \mu_{kijt}(\delta_{kijt}) = E[Z_{kt}(I_{ij})]$ is modeled directly as follows:

$$\mu_{kijt} = \exp(\beta_0 + \underline{x}_{kt}^T(I_{ij}) \underline{\beta}_k^*).$$

The marginal mean μ_{kijt} is the average of the response $Z_{kt}(I_{ij})$ of the farmer population. Zeger et al. (1988) call the $(q \times 1)$ vector $\underline{\beta}_k^*$ the population-averaged (PA) effects, as opposed to the subject-specific effects (SS), $\underline{\beta}_k$. The PA parameters, $\underline{\beta}_k^*$, represent the effect of changes in explanatory variables on this population average. Generally, SS parameters, $\underline{\beta}_k$, and PA parameters, $\underline{\beta}_k^*$, are identical. They differ only in the exceptional case where the explanatory variables in $\underline{x}_{kt}(I_{ij})$ can be represented as a square of a random effects variable or a product of two random effects variables in $\tilde{\underline{x}}_{kt}(I_{ij})$.

In the specified SS model, the individual random effect is introduced only through the random intercept, δ_{kijt} ; hence, the SS and PA models are based on the same set of explanatory variables and all the components of $\underline{\beta}_k^*$ are identical to the corresponding component of $\underline{\beta}_k$. The PA intercept is $\beta_0 = \frac{1}{2} D_{11}$, where D_{11} is the element for the first row of the first column of the matrix \underline{D} (Grömping, 1996).

Individual behavior, δ_{kijt} , has to be taken into account, because to ignore the heterogeneity of behavior in models with aggregated data leads to the positing of incorrect relationships between land use and the associated explanatory variables (see Section 5.3). The random variable δ_{kijt} plays two main roles in our approach: one is to induce overdispersion between aggregated data and the other is to model the spatiotemporal autocorrelation among individual behavior caused by contextual factors.

2.3. The area model

Usually, individual-specific data $\{Z_{kt}(I_{ij}), \underline{x}_{kt}(I_{ij}); j = 1, 2, \dots, M_i; i = 1, 2, \dots, N\}$ are not available. When this is the case, it is only possible to obtain aggregated data at the level of the microunits. The essential characteristic of our approach is to model this aggregated data using an individual model of the choice of $Z_{kt}(I_{ij})$ on the part of the manager of the j th farm and allowing heterogeneity of behavior.

The aggregated data $\underline{Z}_t(A_i) = [Z_{1t}(A_i) Z_{2t}(A_i) \cdots Z_{kt}(A_i) \cdots Z_{Kt}(A_i)]^T$ are modeled as random realizations of a multinomial distribution, $MN(Z(A_i); \underline{p}_{it})$, where \underline{p}_{it} is a $(K \times 1)$ vector the generic component of which, p_{ik} , is the probability that a unit of the land of the territorial microunit A_i is assigned the use $k = 1, 2, \dots, K$ in the period t , and

$Z(A_i) = \sum_{k=1}^K Z_{kt}(A_i)$ is the total units of land (the area of A_i).

In order to simplify the estimation of the model, a multinomial Poisson transformation (Baker, 1994; Chen and Kuo, 2001; Lang, 1996) is used. In this method, the components of $\underline{Z}_t(A_i) = [Z_{1t}(A_i) Z_{2t}(A_i) \cdots Z_{kt}(A_i) \cdots Z_{Kt}(A_i)]^T$ are considered to be independent variables of the Poisson distribution; hence, the distribution of the vector $\underline{Z}_t(A_i)$ is $\text{MN}(Z(A_i); \underline{p}_{it})$, where the generic component of \underline{p}_{it} is $p_{kit} = \frac{\mu_{kit}}{\sum_{k=1}^K \mu_{kit}}$, μ_{kit} is the marginal mean of $Z_{kt}(A_i)$, i.e., $\mu_{kit} = E Z_{kt}(A_i) = \sum_{j=1}^{M_i} E Z_{kt}(I_{ij})$. Hereafter, only $(K - 1)$ categories of land use are considered and the surface assignment to the K th category, which will be called “other” crops and uses, will be defined by subtracting the surface of the modeled $(K - 1)$ categories from the total surface of A_i . Consequently, $\mu_{Kit} = Z(A_i) - \sum_{k=1}^{K-1} \mu_{kit}$ and $p_{Kit} = 1 - \sum_{k=1}^{K-1} p_{kit}$.

The mean μ_{kijt} is modeled using the PA model, $\mu_{kijt} = \exp(\beta_0 + \underline{x}_{kt}^T(I_{ij})\underline{\beta}_k^*)$, and to ensure consistency in aggregation the aggregated marginal mean, $\mu_{kit} = \sum_{j=1}^{M_i} \mu_{kijt}$, is modeled as $\mu_{kit} = \sum_{j=1}^{M_i} \exp(\beta_0 + \underline{x}_{kt}^T(I_{ij})\underline{\beta}_k^*)$. Thus, model parameters, $\underline{\beta}_k^*$, are the same in both the disaggregated and the aggregated models, and consistency in aggregation is ensured in the sense that the system of relationships (between the land use and the explanatory variables) estimated using the aggregated data at microunit level is coherent with the system of relationships specified at the level of the individual farmers. Since individual specific data on $\{\underline{x}_{kt}(I_{ij}); j = 1, 2, \dots, M_i; i = 1, 2, \dots, N\}$ are not available, these will be replaced by the observed average values at the microunit level, $\underline{x}_{kt}(A_i)$, assuming that $\{\underline{x}_{kt}(I_{ij}) = \underline{x}_{kt}(A_i); \forall j = 1, 2, \dots, M_i\}$.

Hence, the model finally adopted for the aggregated data is

$$\mu_{kit} = M_i \exp(\beta_0 + \underline{x}_{kt}^T(A_i)\underline{\beta}_k^*)$$

and

$$p_{kit} = \frac{\exp(\underline{x}_{kt}^T(A_i)\underline{\beta}_k^*)}{\sum_{k=1}^K \exp(\underline{x}_{kt}^T(A_i)\underline{\beta}_k^*)}.$$

Note that p_{kit} is independent of the PA intercept parameter, β_0 , and that the probabilities p_{kit} are uniform when none of the explanatory variables are significant ($\underline{\beta}_k^* = \underline{0}$). In practice, most of the explanatory variables included in $\underline{x}_{kt}(I_{ij})$ are proxies for land rents from alternative use (e.g., input and output prices) and relevant policy variables (Miller and Plantinga, 1999). These variables take the same value at the I_{ij} and A_i levels. However, there are variables, such as land quality measures, that vary among I_{ij} within A_i . In this case, the error in measurement due to replacing these variables by their average value at the A_i level is compounded with the individual behavior effect, δ_{kijt} . Their effects on the model's result are the same as the effect of ignoring individual behavior altogether (see Section 5.3), where aggregated data are used. Note that

in our approach, δ_{kijt} allows us to account for both individual behavior and errors in measuring the explanatory variables. However, in the approach where a random variable, such as δ_{kijt} , is introduced to account solely for errors in measuring the explanatory variables, this variable should be equal to zero when there are no such errors. Then, the individual behavior effects cannot be taken into account. In order to account for the sampling error due to the estimation (aggregation) of the dependent variable, a term of random disturbance, e_{kijt} , will be introduced in the working model used to estimate $\underline{\beta}_k^*$ (see Sections 3 and 4.4).

2.4. Modeling the spatiotemporal correlation structure

The structure of the spatiotemporal autocorrelation is modeled on the assumptions that the aggregated process, $\{Z_{kt}(A_i); i = 1, 2, \dots; t = 1, 2, \dots\}$, is stationary of the second order and isotropic. The covariance matrix of this process, $\underline{V}_k = \text{Var} \underline{Z}_k$, is modeled on the assumption that its diagonal elements are of the form $V Z_{kt}(A_i) = \mu_{kit} + \phi_k^2 \mu_{kit}^2$, where ϕ_k^2 is a nonnegative parameter. The off-diagonal elements are modeled using a correlogram function, $\rho_{kit}(h_i, h_t; \underline{\theta})$. Hence, $\underline{V}_k = \text{Var} \underline{Z}_k = v^{1/2}(\underline{\mu}) \underline{R}(\underline{\theta}) v^{1/2}(\underline{\mu})$, where $v(\underline{\mu}) = \text{diag}(V Z_{kt}(A_i))$ and $\underline{R}(\underline{\theta})$ is a correlation matrix whose components are $\rho_{kit}(h_i, h_t; \underline{\theta})$. To specify this correlogram function, it will be assumed that the temporal and spatial processes can be separated.

The marginal variance of $Z_{kt}(A_i)$ is $V Z_{kt}(A_i) = \sum_{j=1}^{M_i} \mu_{kijt} + V_{\delta_{kijt}} \sum_{j=1}^{M_i} \mu(\delta_{kijt})$, and includes a component, $V_{\delta_{kijt}} \sum_{j=1}^{M_i} \mu(\delta_{kijt})$, that reflects the heterogeneity of behavior among individual farmers. The area models found in the literature are homogeneous, i.e., they assume that $V_{\delta_{kijt}} \mu(\delta_{kijt}) = 0$, with which part of the variability is ignored. However, there is in fact greater variability (overdispersion) than appears in the aggregated data, because of the fact that when aggregating to the level of territorial microunits, the variation between plots, I_{ij} , within the microunits A_i , is eliminated. Simulation (see Section 5.3) is used to evaluate the consequences of ignoring this overdispersion. Here, the variability of $\sum_{j=1}^{M_i} \mu(\delta_{kijt})$ will be modeled on the assumption that its variation coefficient, ϕ_k , is constant so that $V_{\delta_{kijt}} \sum_{j=1}^{M_i} \mu(\delta_{kijt}) = \phi_k^2 \mu_{itk}^2$ and $V Z_{kt}(A_i) = \mu_{kit} + \phi_k^2 \mu_{kit}^2$. Bourlange and Doz (1988) show that this represents a good approximation to a wide set of variance functions. In fact, this is the variance of the negative binomial distribution and can be derived as the variance of a mixed Poisson lognormal model.

Contextual factors shared by farmers who cultivate the land in the same microunit at the same time induce positive autocorrelation (Hart, 1980) between specific individual responses, $\text{Cov}(\mu(\delta_{kijt}), \mu(\delta_{kijt})) > 0$. This is reflected in a spatiotemporal positive autocorrelation of $Z_{kt}(I_{ij})$ and, as a result, also of $Z_{kt}(A_i)$ (Fingleton, 1988). The area models found in the literature ignore the random effect, δ_{kijt} , and, as a result, they ignore this spatiotemporal autocorrelation. Simulation (see Section 5.3) is used to evaluate the consequences of ignoring the spatiotemporal autocorrelation for inference.

2.5. Regional model

At the regional level, $\underline{Z}_t = [Z_{1t} \ Z_{2t} \ \dots \ Z_{kt} \ \dots \ Z_{Kt}]^T$, the surface area occupied by each category of use, $Z_{kt} = \sum_{i=1}^N Z_{kt}(A_i)$, is modeled as a random variable, as is the total of the random variables $Z_{kt}(A_i)$. The distribution characteristics of Z_{kt} are derived from the distribution of $Z_{kt}(A_i)$. In particular, the expected value of Z_{kt} is $\mu_{kt} = \sum_{i=1}^N \mu_{kit}$ and the variance of Z_{kt} is $V(Z_{kt}) = \underline{1}_N^T \text{Var}(\underline{Z}_{kt}) \underline{1}_N$ where $\underline{Z}_{kt} = [Z_{kt}(A_1) \ Z_{kt}(A_2) \ \dots \ Z_{kt}(A_i) \ \dots \ Z_{kt}(A_N)]^T$ and $\underline{1}_N$ is a vector column ($N \times 1$) of ones. In this article, a working model is specified at the level of the microunits for the structure of the average μ_{kit} and the structure of spatiotemporal autocorrelation of the data $\rho_{kit}(h_i, h_t)$. The characteristics of the model at the regional level are derived from those of this working model.

3. Estimation

The marginal distribution of $Z_{kt}(A_i)$ is a mixture of the Poisson distribution and the distribution of δ_{kit} . This mixture of distributions is complex and its likelihood is analytically intractable, because it requires the calculation of high-dimensional integrals (McCulloch and Searle, 2001, p. 263). Full maximum likelihood is not suitable for estimating parameters in these mixed distributions. Hence, quasi-likelihood (presented in McCullagh, 1983 and Wedderburn, 1974) and GEE (presented in Godambe and Thompson, 1989) can be used. These approaches do not require knowledge of the mixture distribution. Desmond (1997) shows how these two areas overlap and complement each other. Only the relationship between the mean and covariates and between the mean and variance of the mixing distribution are required. There is an extensive literature on using quasi-likelihood or GEE for modeling discrete (binary and count data) and continuous data in both longitudinal (Liang and Zeger, 1986; Prentice, 1988; Zeger, 1988; Zeger and Liang, 1986; Zeger et al., 1988) and spatial (Albert and McShane, 1995; McShane et al., 1997) data sets. In these papers, the major focus of interest relates to regression parameters and the precision of their estimates, but they do not treat the prediction problem; neither a predictor nor a prediction error is given. In this article, a predictor is considered as well as its mean squared error. Details of the parameters estimation procedure, the estimation of the spatiotemporal autocorrelation structure, the predictor, and the mean squared error of the predictor are given in the Appendix.

4. Application

In order to assess the effectiveness of the proposed approach to modeling land use, data were obtained from an area of the Lower Guadalquivir (Spain). This area, of about 790,000 hectares, is located between UTM coordinates (212; 4,078) and (394; 4,177) in zone 30 (in km). The climate in the area is

Mediterranean and the yearly averages for rainfall and temperature are 500–800 mm and 15–19°C, respectively. The main crops in this area are sunflower, soft wheat, and sugar beet. For the purposes of the model, these three crops are studied in separate categories and the remaining crops and land uses are grouped in a fourth category called “other.”

4.1. Behavior indicators

According to Ricardo’s theory on income from land (Hardie and Parks, 1997; Polsky and Earsterling, 2001) and Thünen’s on the spatial distribution of crops (Nelson, 2002), it is assumed that the individual choice of the farmers reflects the highest values of land use and that variations in use reflect changes related to the order of the values of alternative uses.

Central to these theories are local prices for outputs and inputs, because these prices influence farmers’ behavior. Unfortunately, we are not able to include these variables because they are not available at the local level (territorial microunits). Instead, EU grant aid (which is officially defined at local level), together with physical medium conditions (evaluated at the territorial microunits level) has been included as behavioral indicators. The value of agricultural land depends, to a large extent, on the conditions of the physical medium (soil and climate). In this article, the conditions of the physical medium in the territorial microunits are evaluated using an agrometeorological model and used as variables that affect farmers’ behavior. Agrometeorological models have been widely used for optimizing crop handling systems (Boote et al., 1996; Whisler et al., 1986) and, in the last few years, their use for evaluating the effects of climatic change on agriculture has emerged as a recognizable field of research (Adams et al., 1990, 1995; Pickering et al., 1995). We use the agrometeorological model Crop Growth Monitoring System (CGMS) (van Raaij and van der Wal, 1994).

In CGMS, the conditions of the physical environment are evaluated using three sources: soil map data, data from meteorological stations, and a model of the vegetative development for each crop. CGMS has been calibrated in our work area and the following indicators of the quality of the physical medium have been obtained for each territorial microunit: (i) the weighted biomass and the weighted dry matter in storage organs, (ii) leaf area index, (iii) development stage, and (iv) water consumption and requirement. The dry matter in the storage organs shows the greatest linear correlation with the yield of the three crops that were considered. Hence, it is used here as an indicator of quality of the physical medium for sunflower (SF_CGMS), soft wheat (SW_CGMS), and sugar beet (SB_CGMS). For behavior indicators, we use EU direct grant aid (€/ha) for sunflower (SF_AID) and soft wheat (SW_AID), as defined in the Mac Sharry reform of the Common Agricultural Policy.

In order to evaluate the dynamics of land use, (land) uses of the year $t - 1$ are included as explanatory variables of the uses of year t . The crop area for one crop in year $t - 1$ is included as an explanatory variable for the area of this and the other

crops in the year t . More precisely, the sunflower crop area in year $t - 1$ is included as an explanatory variable of the area of this crop (SF_1_SF) and of the area of soft wheat (SW_1_SF) and sugar beet (SB_1_SF), in year t . In the same way, the following variables are included: (SW_1_SW), (SF_1_SW), and (SB_1_SW), for soft wheat; (SB_1_SB), (SF_1_SB), and (SW_1_SB) for sugar beet; and (SF_1_OT), (SW_1_OT), and (SB_1_OT) for crops in the category “other.”

4.2. The data

Land use data were obtained using a spatial sampling scheme proposed by the EU Joint Research Centre (Gallego et al., 1994). In order to carry out the sampling, topographical maps were used. The sampling unit was the UTM 1 km by 1 km grid (in Spain this is conventionally reduced to 700 m by 700 m). These sampling units were grouped into blocks of 10 km by 10 km ($N = 127$). A systematic sample of three sample units per block was selected (Ambrosio et al., 2003), so that the sample size was 381 sampling units (127×3). In each one of the selected sampling units, data on land use were observed during the $T = 8$ years from 1995 to 2002 [ESYRCE (1995–2002)]. A total of 3,048 observations were therefore used (381×8) to fit the model. In the data set for each panel, land use varies among sampling units and from year to year within the same sampling unit. Only data for sunflower, soft wheat, and sugar beet are considered: the average per sampling unit is 8.72 hectares for sunflower, 0.85 hectares for soft wheat, and 0.78 hectares for sugar beet.

In order to explain the variability in land use, data on EU grant aid and on physical conditions are used. EU grant aid for sunflower and soft wheat (there is no grant aid for sugar beet) is defined for each crop at the municipality level, as a function of crop yield. In 1995, this ranged from 259 to 532 €/ha for sunflower and from 103 to 212 €/ha for soft wheat. In 2002, it ranged from 130 to 268 €/ha for sunflower and from 113 to 233 €/ha for soft wheat.

Physical conditions were assessed by evaluating the dry matter in the storage organs of each crop, using the agrometeorological CGMS model. The dry matter in the storage organs for sunflower (SF_CGMS), soft wheat (SW_CGMS), and sugar beet (SB_CGMS) were evaluated within each sampling unit, in each one of the smallest cartographable units registered on the land maps. SF_CGMS ranges among cartographable units and years from 1,850 to 6,467 kg/ha, SW_CGMS from 4,963 to 15,593 kg/ha, and SB_CGMS from 9,752 to 24,126 kg/ha.

4.3. Data aggregation

The 10 km by 10 km block was used as the territorial microunit, A_i , and all data were aggregated to this level. The area of each category of use was estimated in each microunit using the sample average of the values observed in the three sampling units. EU grant aid was aggregated at the level of

each territorial microunit as a weighted average, using the municipal area within the block as a weighting coefficient. The CGMS results have the smallest cartographable unit registered on the land maps. These were aggregated to the level of the territorial microunits as a weighted average using the area of the cartographable unit within the block as a weighting coefficient.

4.4. Working model

The working model consisted of specifying at the level of the territorial microunits: (i) the structure of the mean, μ_{kit} , and (ii) the structure of the spatiotemporal autocorrelation, V_k :

- (i) $\eta_{kit} = \log(\mu_{kit}) = \beta_0 + \mathbf{x}_{kit}^T \underline{\beta}_k$
- (ii) $\underline{V}_k = \text{Var} \{ \hat{Z}_{kit}; i = 1, 2, \dots, N; t = 1, 2, \dots, T_i \}$
 $= [\Sigma_e + v^{1/2}(\underline{\mu}_k)] R(\underline{\theta}_k) [\Sigma_e + v^{1/2}(\underline{\mu}_k)]$.

Lee and Zhao (1997) show that the link function, $g(\cdot)$, does not affect the results. The option used here is that most commonly used: the logarithmic, $\log(\cdot)$.

Data related to land use were estimated using as estimator the sample average, $\hat{Z}_{ks,t}$, so that $\hat{Z}_{kit} = Z_{kt}(A_i) + e_{kit}$, where e_{kit} is the sampling error, of zero mean and variance $\sigma_{e_{ks,t}}^2$. In the model, the error was taken into account in the data through $\Sigma_e = \text{diag}\{\sigma_{e_{ks,t}}^2\}$. In the cases where $Z_{kt}(A_i)$ was observed without error, $\Sigma_e = \underline{0}$.

The structure of spatiotemporal autocorrelation was introduced in the model through the matrix $R(\underline{\theta}_k)$, whose $[NT \times NT]$ elements were calculated on the basis of the correlogram function, $\rho_{kit}(h_i, h_t; \underline{\theta}_k)$, where $\underline{\theta}_k$ was the vector of the parameters used to specify that structure.

The estimator $\underline{\hat{\beta}}$ of $\underline{\beta} = [\beta_0 \ \underline{\beta}_1 \ \dots \ \underline{\beta}_k \ \dots \ \underline{\beta}_{K-1}]^T$ is consistent if the mean structure, μ_{kit} , and the variance function, $v(\mu_{kit})$, are specified correctly, although the function of autocorrelation, $\rho_{kit}(h_i, h_t)$, used for the estimation is only approximate (Liang and Zeger, 1986; Zeger and Liang, 1986). This is important because $\rho_{kit}(h_i, h_t)$ changes with the size and the form of the microunits (Gotway and Young, 2002), and the estimator of $\underline{\beta}_k$ is consistent regardless of that size and form. In dynamic models, the conditions at the start must be specified. In this study, the starting conditions were those observed in the first years observed in the sample (Hsiao, 2003).

5. Results

Fig. 2 shows the empirical semivariogram of data residues of land use, after discounting the effect of the explanatory variables. As we can see, the variability of those residues grows (or diminishes) to the same extent as the spatiotemporal distance between observations. This tendency is attributable to the unobserved contextual factors that place the farmers in the

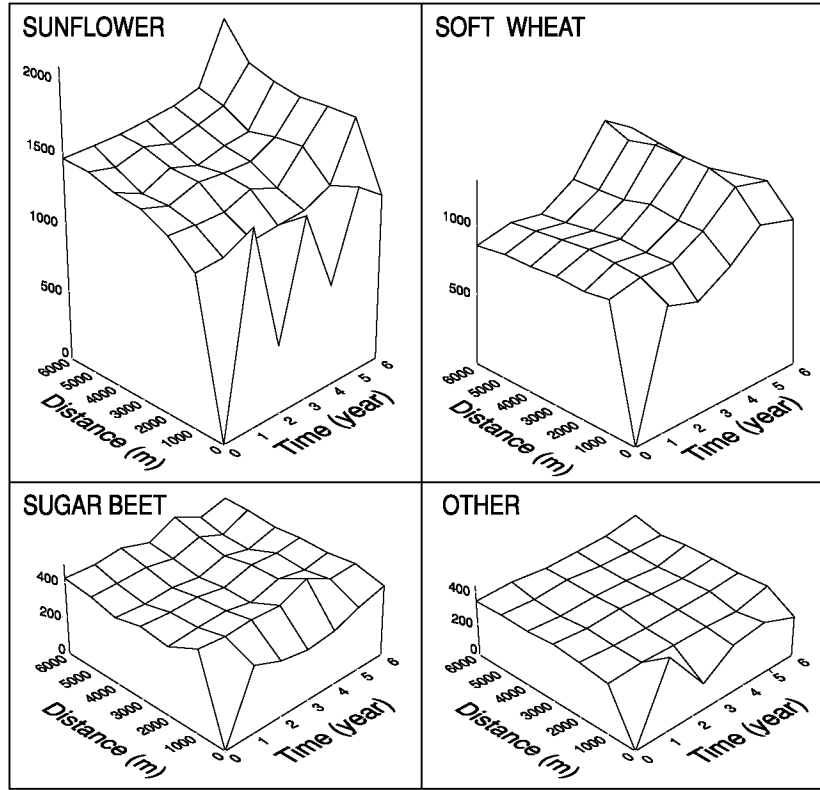


Fig. 2. Empirical spatiotemporal semivariograms.

same spatiotemporal context (they cultivate plots close to each other at the same time), exhibiting similar (though heterogeneous) behavior and, therefore, little variability is observed in their decisions. The difference in behavior and therefore the variability of decisions increases as the spatiotemporal distance increases.

5.1. Estimations of the coefficients and of the power stretch

The estimations of the coefficients β are shown in Table 1, including the standard errors in brackets. The significance level that corresponds to the asymptotic t -ratio is marked with an asterisk. Elasticity estimations of each category of use are also included with regard to a sample average of the behavior indicators: mean elasticity was calculated as the product of the estimated value for the coefficient and the mean of the relevant explanatory value.

The local physical environment significantly influences the farmers' decisions about land use: the indicator sign of these conditions is positive for sugar beet (SB_CGMS) and negative for sunflower (SF_CGMS) and soft wheat (SW_CGMS). Consequently, the land area assigned to sugar beet increases as the conditions of the physical medium improve, to the detriment of soft wheat and sunflower crops. Sugar beet appears to be more elastic to changes in the conditions of the physical medium (0.74) than soft wheat (−0.025) and sunflower (−0.02).

EU grant aid for sunflower crops (SF_AID) and soft wheat (SW_AID) significantly influence farmers' decisions about land use. The area dedicated to either of these two crops is elastic with regard to this factor. The elasticity is positive and greater for soft wheat (2.18) than for sunflower (1.61).

Decisions taken in the past about land use have a noticeable effect on the present and future behavior of the farmers. This is shown by crop rotation. Sugar beet is the only one of the three crops considered whose area shows a positive association with itself from one year to the next: the coefficient sign of the variable SB_1_SB is positive, so that an increase (a decrease) of sugar beet surface in the year $(t - 1)$ tends to remain in the year t , although at a low rate (elasticity of 0.02). However, the sunflower (SF_1_SF) and soft wheat (SW_1_SW) areas show negative elasticities with regard to the areas of the previous year: an increase (a decrease) of the sunflower area in the year $(t - 1)$ tends to make the sunflower area decrease (increase) in the year t and the same applies to soft wheat. The elasticity is higher in sunflower (−0.30) than in soft wheat (−0.12).

The elasticity of an area occupied by one crop compared with another from one year to the next is negative. For example, the elasticity of the sunflower area in the year t compared to the changes of the soft wheat surface in the year $(t - 1)$ (SF_1_SW) is negative (−0.08), so that an increase (decrease) of the soft wheat area in the year $(t - 1)$ tends to result in a decrease (increase) in the sunflower area in the year t . The relation between the system formed by the three crops considered

Table 1
Estimation of the coefficients and of the elasticities

Variable	Estimated coefficients (standard deviation)	Average elasticity
Intercept	9.1** (0.44099)	9.1
SF_CGMS	-0.00001** (3.6E-06)	-0.02
SW_CGMS	-0.0001** (3.4E-05)	-0.025
SB_CGMS	0.0007** (0.00013)	0.74
SF_AID	0.0054** (0.00036)	1.61
SW_AID	0.0125* (0.00505)	2.18
SF_1_SF	-0.0002** (6.3E-05)	-0.30
SW_1_SW	-0.0003** (0.0001)	-0.12
SB_1_SB	0.0001* (4.8E-05)	0.02
SF_1_SW	-0.0002** (3.4E-05)	-0.08
SF_1_SB	-0.0005** (8.2E-05)	-0.1
SW_1_SF	-0.0004** (0.00011)	-0.6
SW_1_SB	-0.0004* (0.00016)	-0.08
SB_1_SW	-0.0003* (0.00013)	-0.12
SB_1_SF	-0.0004** (8.9E-05)	-0.6
SF_1_OT	-0.0003** (4.6E-05)	-2.37
SW_1_OT	-0.0007** (0.00012)	-5.53
SB_1_OT	-0.0005** (6.0E-05)	-3.95

* and ** denote significance at 5% and 1%, respectively.

(sunflower, soft wheat, and sugar beet) and the remaining crops and categories of land use is established through the category “others.” The elasticity of an area occupied by each of the three crops considered with regard to the area occupied by “others” is negative. An increase (decrease) of the surface of “others” in the year $t - 1$ tends to result in a decrease (increase) in the surface of each of the three crops considered in the year t . Yet again, the soft wheat surface is more elastic (-5.53) than that of sugar beet (-3.95) and sunflower (-2.37).

5.2. Forecasts at the level of territorial microunits

To forecast the land use in each microunit, using (A.6) (see Appendix), we predicted both the choice of the average farmer and the outcome of the interactions of individual farmers with contextual factors. This outcome was estimated as a weighted

average of the residues observed in neighbouring microunits. The weight of a microunit in the forecast of that outcome is proportional to the value of the variogram function, so that it increases (decreases) when the spatiotemporal distance between them decreases (increases). This is an exact predictor in the sense that the values observed in the sample are forecast without error, and an optimum predictor in the sense that the values not observed are forecast with minimum error.

A scenario was considered in which EU grant aid for crops was removed entirely and the forecasts of the areas of each of the crops considered were obtained on the assumption that the conditions of the physical medium were the same as those observed in the sample. The impact on the average area per microunit is shown in Figs. 3 and 4: the sunflower and soft wheat areas would decrease, while those of the sugar beet and of “other” crops and land uses would increase. As shown in Fig. 3, the interannual variability of the forecasts (which is attributable to the variability in the climatic conditions and unobservable contextual factors) is adjusted to that which is effectively observed. This is so except in the case of sugar beet, whose interannual variability increases in the conditions of the scenario due to the high elasticity compared with the area of “other” crops and land uses, with which it competes for the land free from sunflower and soft wheat.

The spatial distribution of the impact is shown in the map of Fig. 4, which was built on the basis of the crop area forecasts considered in each microunit and under different scenarios. It can be observed that the area free from sunflower and soft wheat is partly taken over by sugar beet and partly by “other” crops and land uses.

5.3. Effect of ignoring the heterogeneity of individual behavior and the spatiotemporal autocorrelation: simulations

In order to evaluate the proposed model, the results were compared with the homogenous and independent model that resulted from ignoring overdispersion and autocorrelation, while maintaining the same specification of the mean structure. The estimator of the parameters was consistent in both cases, which is why the comparison is limited to the variance of the estimator of the parameters and that of the predictor. A hundred panel samples of the same size as that observed ($N = 127$ territorial microunits repeated over a period of $T = 8$ years) were simulated on the basis of the fitted working model (i.e., using the coefficient of Table 1 and the semivariograms of Fig. 2).

Both models were fitted to each of the hundred samples and the deviation of the variance of the estimator was evaluated with regard to that of the “true” generator model of the samples. $\hat{\text{Var}}(\hat{\beta}_l)$ was calculated using (A.2) (see Appendix) and the l th simulated sample ($l = 1, 2, \dots, L$). $\hat{V}_{kl}(\hat{Z}_{kt}(A_l) - Z_{kt}(A_l))$ was calculated using (A.7) (see Appendix). The bias of the estimator of the covariance matrix is $B(\hat{\text{Var}}(\hat{\beta})) = \hat{\text{Var}}(\hat{\beta}) - \underline{\Lambda}$, where $\hat{\text{Var}}(\hat{\beta}) = \frac{1}{L} \sum_{l=1}^L \hat{\text{Var}}(\hat{\beta}_l)$ and $\underline{\Lambda} =$

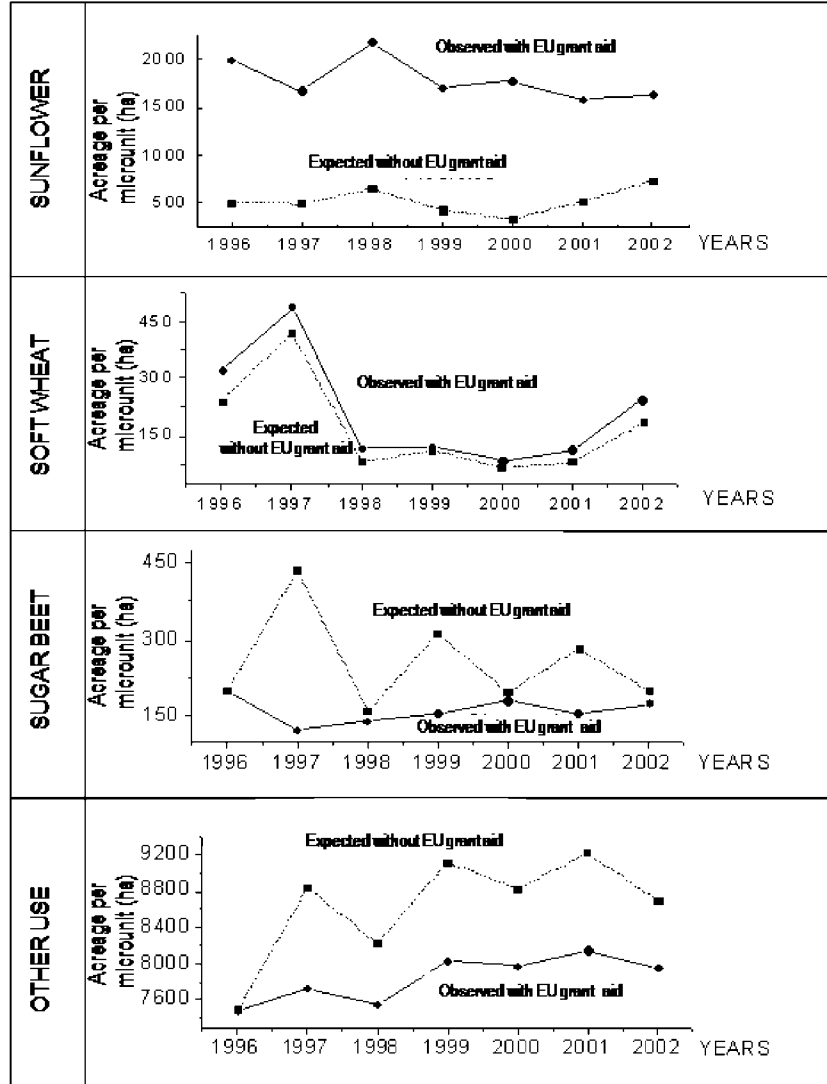


Fig. 3. Interannual variability of the forecast of the impact of suppressing EU grant aid on crop acreage per microunit.

$\frac{1}{L} \sum_{l=1}^L (\hat{\beta}_l - \beta)(\hat{\beta}_l - \beta)^T$. The bias of the estimator of the forecast error was $B(\hat{V}_k(\hat{Z}_{kt}(A_l) - Z_{kt}(A_l))) = \hat{V}_k(\hat{Z}_{kt}(A_l) - Z_{kt}(A_l)) - V_k(\hat{Z}_{kt}(A_l) - Z_{kt}(A_l))$, where $\hat{V}_k(\hat{Z}_{kt}(A_l) - Z_{kt}(A_l)) = \frac{1}{L} \sum_{l=1}^L \hat{V}_{kl}(\hat{Z}_{kt}(A_l) - Z_{kt}(A_l))$ and $V_k(\hat{Z}_{kt}(A_l) - Z_{kt}(A_l))$ was calculated using (A.7) (see Appendix) and the true values of the parameters. The effect of ignoring the overdispersion and the spatiotemporal autocorrelation was evaluated by comparing the bias when the overdispersion and the spatiotemporal autocorrelation were taken into account and when they were ignored assuming $V_{\delta_{itk}} \mu(\delta_{itk}) = 0$ and $\rho_{kit}(h_i, h_t; \underline{\theta}_k) = 0$, that is $R(\underline{\theta}_k) = \underline{I}$ for $k = 1, 2, \dots, K$.

It was found that both models underestimated the standard error of the estimator of the parameters and that the degree of underestimation was noticeably higher in the homogenous and independent model: on average, 38.5% and 11.3%, respectively.

(These results are consistent with those registered by other studies using the same methodology, although only with temporal data [Zeger, 1988] or only with spatial data [McShane et al., 1997]). Consequently, both the amplitude of the reliability intervals, and the error (of the first species) of the significance test were underestimated. The power of the test for the hypothesis that a certain coefficient is nil and the alternative that this coefficient is equal to the value of Table 1 was calculated. That power was, on average, 68.9% with the proposed model and fell to 38.8% with the homogenous and independent model. This suggests a substantial fall in the capacity of this latter model to detect significant indicators of the farmers' behavior: the probability that an effectively significant coefficient will be detected is reduced to half with the homogenous and independent models. In the case of coefficients associated with the dynamic component of the model, those probabilities are reduced still

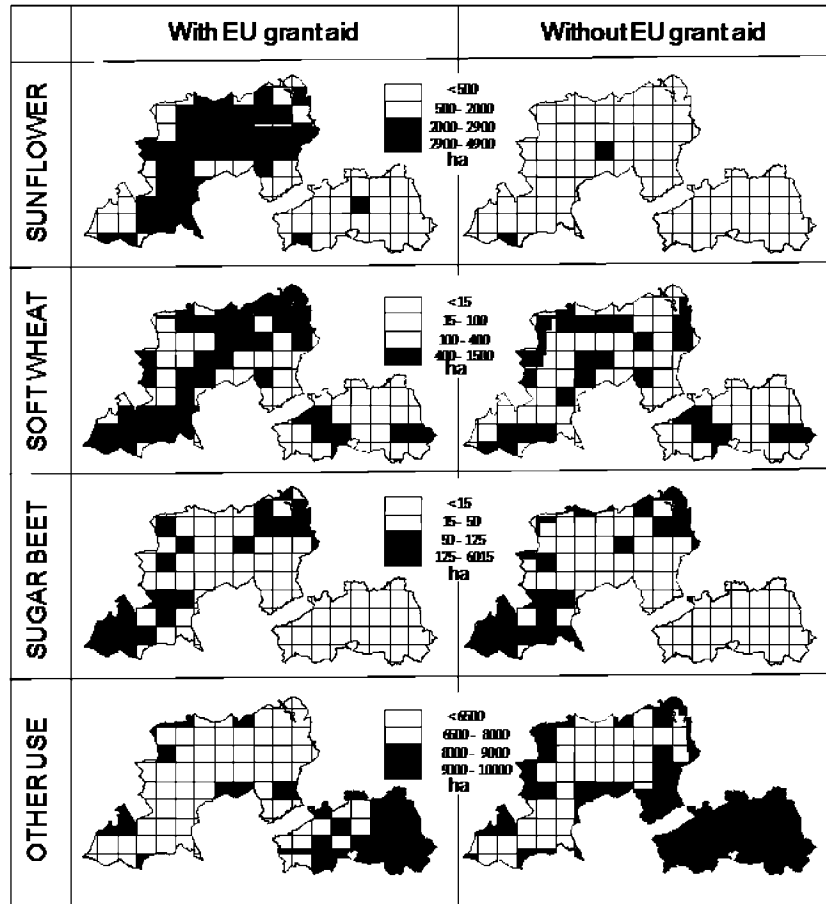


Fig. 4. Mapping the impact of suppressing EU grant aid on crop acreage per microunit in the studied zone.

further, so that the probability of detecting true state dependence with the homogenous and independent model is very low (average power 16.6%).

It was found that both models underestimate the forecast error and that the degree of underestimation is remarkably higher with the homogeneous and independent model. With the proposed model, in half the simulated samples the underestimation increases from 4.7 to 24.9%, while with the homogenous and independent model it increases from 23.7 to 47.3%. The underestimation was less than 25% in three out of every four samples simulated with the proposed model, while with the homogenous and independent model the underestimation was less than 25% in only one out of every four samples. Consequently, with the homogenous and independent model, the amplitude of the reliability intervals of the forecasts is underestimated and the underestimation increases remarkably.

6. Discussion

A dynamic area model was proposed for analyzing changes of agricultural land use that takes into account the heterogeneity of behavior of individual farmers. When area mean data are

used, that heterogeneity is not observed and its effects must be modeled. An approach based on latent variables was proposed to model those effects. In practice, the statistical distribution of those latent variables is unknown. A clear advantage of the suggested approach is that it is not necessary to know this distribution: it is enough to model the structure of the mean and spatiotemporal autocorrelation. In the dynamic models, the initial conditions must be specified. In this research, those conditions were taken to be those observed in the first year of the sample, as if the process of changing land use began at that moment. This is a weakness of the proposed model and further research will be required to improve this feature.

The model was applied to an area of the Lower Guadalquivir (Spain). An agrometeorological model was used to evaluate the conditions of the physical medium (soil and climate). It was found that the influence of these conditions on the decisions of farmers at the time of choosing their alternative crop is statistically significant. At the same time, it was found that past experience has a genuine effect on the present and future behavior of the farmer with respect to land use. Consequently, static models of land use can encounter problems that are caused by the omission of relevant variables. EU grant

aid for the cultivation of sunflowers and wheat also has a statistically significant influence on the choice of crop alternative for the farmers concerned.

Our approach to analyzing changes of land use can contribute to the improvement of the analysis of agricultural policies and the evaluation of their impact in three principal ways. One, the model integrates individual farmer behavior. Two, the dynamic nature of the model allows the temporal distribution of the impact of the policy measures to be monitored. Three, the model is spatially explicit and it allows the forecasting of future changes of land use and the evaluation of environmental effects on natural resources associated with these changes with the required spatial resolution.

Together with GIS tools, the proposed model enables the generation of impact maps of policies on land use. Together with environmental indicators, it enables the assessment of the environmental impacts of agricultural policy, with the required spatial resolution. Thus, it can also contribute to the promotion of agricultural policy in a more precise way.

However, the model proposed requires a greater amount of data to be processed in order to take overdispersion and spatiotemporal autocorrelation into account. Nevertheless, the additional effort involved in this increase in data processing appears to be worthwhile. This is because of the fact that underestimation of errors in the model parameters' estimator and the predictor is substantially reduced. Consequently, the reliability intervals and the tests of significance are more adjusted to reality than those obtained on the basis of models in which the overdispersion is ignored and independence is assumed.

Acknowledgments

Financial support from the Spanish Ministry for Education and Science for Project CGL2005-02589/CLI and from the Junta de Andalucía is gratefully acknowledged. We wish to thank L. A. Rubio, J. Montañés, L. Balairón, J. M. Martín, S. E. Petisco, A. Garrido, and E. Jiménez for their help in the case study. The referees and the co-editors-in-chief have contributed to improving this article, making helpful comments and suggestions.

Appendix

A.1. Model parameters estimation

Let $g(\underline{\mu}) = \underline{X}\underline{\beta}$ be the model specified for the whole of the $NT(K-1)$ data of the sample, where $g(\cdot) = \log(\cdot)$, $\underline{\beta} = [\beta_0 \ \underline{\beta}_1 \ \cdots \ \underline{\beta}_k \ \cdots \ \underline{\beta}_{K-1}]^T$ is a $(1 + (K-1)q) \times 1$ vector with $\underline{\beta}_k = [\beta_{k1} \ \beta_{k2} \ \cdots \ \beta_{kq}]^T$, and $\underline{X} = [\underline{1} \ \underline{\tilde{X}}]$ where $\underline{\tilde{X}}$ is a diagonal matrix by blocks in accordance with $\underline{\beta}$ whose k th row is \underline{x}_{kit}^T . For the estimator $\hat{\underline{\beta}}$ of $\underline{\beta}$ to verify the restriction $\sum_{k=1}^K \mu_{kit} = \sum_{k=1}^K Z_{kt}(A_i) = Z_t(A_i)$, it is necessary for the matrix \underline{X} to

include a vector column of ones, $\underline{1}$, with which the parameter β_0 is associated (Christensen, 1990, p. 189). The k th row of \underline{X} has a one as the first element and the remaining elements are nil, except the elements associated with $\underline{\beta}_k$, which are \underline{x}_{kit} . The estimator $\hat{\underline{\beta}}$ is obtained by solving the equation system

$$\underline{X}^T \underline{W} \underline{X} \underline{\beta} = \underline{X}^T \underline{W} \underline{Z} \quad (\text{A.1})$$

where $\underline{W} = \underline{D}^T \underline{V}^{-1} \underline{D}$, $\underline{V} = \text{diag}\{\underline{V}_k\}$, $\underline{D} = \text{diag}\{\underline{\mu}_k\}$, $\underline{\mu}_k = [\mu_{k1} \ \mu_{k2} \ \cdots \ \mu_{kt} \ \cdots \ \mu_{kT}]^T$, $\underline{\mu}_{kt} = [\mu_{k1t} \ \mu_{k2t} \ \cdots \ \mu_{kit} \ \cdots \ \mu_{kNt}]^T$, $\underline{Z} = \underline{\eta} + \underline{D}^{-1}(\underline{Z} - \underline{\mu})$, $\underline{Z} = [\underline{Z}_1 \ \underline{Z}_2 \ \cdots \ \underline{Z}_k \ \cdots \ \underline{Z}_K]^T$, $\underline{Z}_k = [\underline{Z}_{k1} \ \underline{Z}_{k2} \ \cdots \ \underline{Z}_{kt} \ \cdots \ \underline{Z}_{kT}]$, and $\underline{Z}_{kt} = [Z_{kt}(A_1) \ Z_{kt}(A_2) \ \cdots \ Z_{kt}(A_i) \ \cdots \ Z_{kt}(A_N)]^T$. The covariance matrix is

$$\text{Var} \hat{\underline{\beta}} = (\underline{X}^T \underline{W} \underline{X})^{-1} (\underline{X}^T \underline{W} \underline{X}) (\underline{X}^T \underline{W} \underline{X})^{-1} \quad (\text{Valliant, 1986}), \quad (\text{A.2})$$

where $\underline{W} = \underline{D}^T \underline{V}^{-1} \text{Var}(\underline{Z}) \underline{V}^{-1} \underline{D}$, and $\text{Var}(\underline{Z})$ is consistently estimated by $(\underline{Z} - \underline{\mu})(\underline{Z} - \underline{\mu})^T$, where \underline{Z} is a vector $(NT(K-1) \times 1)$ whose i th element is $Z_{kt}(A_{S_i})$, $\underline{\mu}$ is replaced by its consistent estimations, and ϕ_k^2 is estimated by $\hat{\phi}_k^2 = \frac{1}{NT-(p+K)} \sum_{i=1}^N \sum_{t=1}^T \frac{[(Z_{kit} - \hat{\mu}_{kit})^2 - \hat{\mu}_{kit}]}{\hat{\mu}_{kit}}$, where $\hat{\mu}_{kit} = \exp(\hat{\beta}_0 + \underline{x}_{kit}^T \hat{\underline{\beta}}_k)$ (Gourieroux et al., 1984). If the overdispersion is ignored ($V_{\delta_{itk}} \mu(\delta_{itk}) = 0$) and the spatiotemporal autocorrelation is ignored ($R(\underline{\theta}_k) = \underline{I}$), then the estimator of $\underline{\beta}$ in (A.1) is reduced to that of the maximum likelihood and (A.2) is reduced to the usual covariance matrix,

$$\text{Var} \hat{\underline{\beta}} = (\underline{X}^T \underline{V}^{-1} \underline{X})^{-1}.$$

For computing $\hat{\underline{\beta}}$ and $\text{Var} \hat{\underline{\beta}}$, a program for the IML procedure of SAS (1997) was written.

A.2. Estimation of the spatiotemporal autocorrelation structure

The correlogram function is $\rho_{kit}(h_i, h_t; \underline{\theta}) = C_{kit}(h_i, h_t) / \sigma_{kit}^2 = 1 - \gamma_{kit}(h_i, h_t) / \sigma_{kit}^2$ where $C_{kit}(h_i, h_t) = \text{Cov}(Z_{kt}(A_i), Z_{kt'}(A_{i'})) = E[(Z_{kt}(A_i) - \mu_{kit})(Z_{kt'}(A_{i'}) - \mu_{kit'}(A_{i'}))]; \forall i, i' | \text{dist}(i, i') = h_i; \forall t, t' | |t - t'| = h_t$ where E denotes the expectation of the distribution of $Z_{kt}(A_i)$ and $\gamma_{kit}(h_i, h_t) = \frac{1}{2} \text{Var}(Z_{kt}(A_i) - Z_{kt'}(A_{i'})) = \gamma_{kit}(\text{dist}(i, i'), |t - t'|); \forall i, i' | \text{dist}(i, i') = h_i; \forall t, t' | |t - t'| = h_t$ is the semivariogram function. It is assumed that $C_{kit}(h_i, h_t)$ is continuous in space and in time, $\lim_{h_i \rightarrow \infty} \lim_{h_t \rightarrow \infty} C_{kit}(h_i, h_t) = 0$ and $\sigma_{kit}^2 = \text{sill} \gamma_{kit}(h_i, h_t) = \lim_{h_i \rightarrow \infty} \lim_{h_t \rightarrow \infty} \gamma_{kit}(h_i, h_t)$. It is assumed that the spatial and temporal processes are separable (De Cesare et al., 2001; De Iaco et al., 2002):

$$\gamma_{kit}(h_i, h_t) = \gamma_{ki}(h_i, 0) + \gamma_{kt}(0, h_t) - \kappa \gamma_{ki}(h_i, 0) \gamma_{kt}(0, h_t),$$

where

$$\kappa = [\text{sill} \gamma_{kit}(h_s, 0) + \text{sill} \gamma_{kit}(0, h_t) - \text{sill} \gamma_{kit}(h_i, h_t)] / \text{sill} \gamma_{kit}(h_i, 0) \text{sill} \gamma_{kit}(0, h_t).$$

In order for $\gamma_{kit}(h_i, h_t)$ to be a valid semivariogram function, κ must be $0 < \kappa \leq 1/\max\{\text{sill}\gamma_{kit}(h_i, 0); \text{sill}\gamma_{kit}(0, h_t)\}$. The spatial component estimator is (Zimmerman and Zimmerman, 1991):

$$\hat{\gamma}_{kit}(h_i, 0) = \frac{1}{2} \frac{1}{|N(h_i, 0)|} \sum_{N(h_i, 0)} (\hat{\epsilon}_{kit} - \hat{\epsilon}_{ki't'})^2 \quad (\text{A.3})$$

where $\hat{\epsilon}_{kit} = \frac{Z_{kit}(A_i) - \hat{\mu}_{kit}}{\sqrt{v(\hat{\mu}_{kit}, \hat{\phi}_k^2)}}$ and $N(h_i, 0) = \{(it, i't') | \text{dist}(i, i') = h_i; t = t'\}$ with $v(\hat{\mu}_{kit}, \hat{\phi}_k^2) = \hat{\mu}_{kit} + \hat{\phi}_k^2 \hat{\mu}_{kit}^2$, and the same for i' . The temporal component estimator is

$$\hat{\gamma}_{kit}(0, h_t) = \frac{1}{2} \frac{1}{|N(0, h_t)|} \sum_{N(0, h_t)} (\hat{\epsilon}_{kit} - \hat{\epsilon}_{ki't'})^2 \quad (\text{A.4})$$

where $N(0, h_t) = \{(it, i't') | |t - t'| = h_t; i = i'\}$. The spatiotemporal semivariogram is estimated by (De Iaco et al., 2001)

$$\hat{\gamma}_{kit}(h_i, h_t) = \frac{1}{2} \frac{1}{|N(h_i, h_t)|} \sum_{N(h_i, h_t)} (\hat{\epsilon}_{kit} - \hat{\epsilon}_{ki't'})^2, \quad (\text{A.5})$$

where $N(h_i, h_t) = \{(i, i'); (t, t') | \text{dist}(i, i') = h_i, |t - t'| = h_t\}$. The global sill $\gamma_{it}(h_i, h_t)$ is estimated graphically, by drawing the empiric semivariogram. As soon as $\text{sill}\gamma_{ki}(h_i)$, $\text{sill}\gamma_{kt}(h_t)$, and $\text{sill}\gamma_{kit}(h_i, h_t)$ have been estimated, they are used to estimate κ , and it is verified that κ satisfies the condition required for the matrix to be positively semidefinite. The spatiotemporal autocorrelation structure is estimated on the basis of (A.3), (A.4), and (A.5), using the Pearson residues of the fitted model assuming $R(\theta_k) = \underline{I}$.

A program for the IML procedure of SAS (1997) was written for computation.

A.3. The predictor and its mean squared error

The unbiased and optimal predictor of $Z_{k,T+1}(A_i)$ (if $\underline{\beta}$, $\phi_{(c)}^2$, and $\underline{\theta}_{(c)}$ are known) is

$$\hat{Z}_{ki,T+1} = \hat{\mu}_{ki,T+1} + \underline{C}_{k(i,i'),(t,T+1)}^T \underline{V}_k^{-1} (\underline{Z}_k - \hat{\underline{\mu}}_k) \quad (\text{Valliant, 1985}) \quad (\text{A.6})$$

where $\hat{\mu}_{ki,T+1} = \exp(\tilde{x}_{ki,T+1} \hat{\underline{\beta}}_k)$ and $\underline{C}_{k(i,i'),(t,T+1)}^T$ is the correlation vector between $Z_{k,T+1}(A_i)$ and each variable in $\{Z_{k,t}(A_i); i = 1, 2, \dots, N; t = 1, 2, \dots, T\}$. If independence is assumed, $\underline{C}_{k(i,i'),(t,T+1)}^T = \underline{0}$, and $\hat{Z}_{ki,T+1}$ it is reduced to $\hat{\mu}_{ki,T+1}$. The mean squared error of the predictor $\hat{Z}_{ki,T+1}$ is estimated by (Valliant, 1985; Vijapurkar and Gotway, 2001),

$$\begin{aligned} V(\hat{Z}_{k,T+1}(A_i) - Z_{k,T+1}(A_{s_i})) &\cong V Z_{k,T+1}(A_{s_i}) \\ &- \underline{C}_{k(i,i'),(t,T+1)}^T \underline{V}_k^{-1} \underline{C}_{k(i,i'),(t,T+1)} + [\mu_{ki,T+1} \tilde{x}_{ki,T+1} \\ &- \underline{C}_{k(i,i'),(t,T+1)}^T \underline{V}_k^{-1} \underline{D} \underline{X}] \text{Var}(\hat{\underline{\beta}}_k) [\mu_{ki,T+1} \tilde{x}_{ki,T+1} \\ &- \underline{C}_{k(i,i'),(t,T+1)}^T \underline{V}_k^{-1} \underline{D} \underline{X}]^T, \end{aligned} \quad (\text{A.7})$$

where

$$V Z_{k,T+1} = \mu_{ki,T+1} + \phi_k^2 \mu_{ki,T+1}^2.$$

A program for the IML procedure of SAS (1997) was written for computation.

[Correction added after online publication 8 February 2008: in (A.7) A_i was changed to A_{s_i}]

References

- Adams, R. M., Rosenzweig, C., Peart, R. M., Ritchie, J. T., McCarl, B. A., Glyer, J. D., Curry, R. B., Jones, J. W., Boote, K. J., Allen, L. H. Jr., 1990. Global climate change and U.S. agriculture. *Nature* 345, 219–224.
- Adams, R. M., McCarl, B. A., Dudek, D. J., Glyer, J. D., 1995. A reassessment of the economic effects of global climate change on U.S. agriculture. *Clim. Change* 30, 147–167.
- Albert, P. S., McShane, L. M., 1995. A generalized estimating equations approach for spatially correlated binary data: applications to the analysis of neuroimaging data. *Biometrics* 51, 627–638.
- Alig, R. J., 1986. Econometric analysis of the factors influencing forest acreage trends in the South-east. *For. Sci.* 32, 119–134.
- Ambrosio, L., Iglesias, L., Marín, C., 2003. Systematic simple design for the estimation of spatial means. *Environmetrics* 14, 45–61.
- Baker, S. G., 1994. The multinomial-Poisson transformation. *Statistician* 43, 495–504.
- Baltzer, H., Braun, P. W., Köhler, W., 1998. Cellular automata models for vegetation dynamics. *Ecol. Models* 107, 113–125.
- Boote, K. J., Jones, J. W., Pickering, N. B., 1996. Potential uses and limitations of crop models. *Agron. J.* 88, 704–716.
- Bourlange, D., Doz, C., 1988. Pseudo-maximum de vraisemblance: expériences de simulation dans le cadre d'un model de Poisson. *Annal. Econ. Stat.* 10, 139–176.
- Chen, Z., Kuo, L., 2001. A note on the estimation of the multinomial logit model with random effects. *Am. Stat.* 55, 89–95.
- Christensen, R., 1990. *Log-Linear Models*. Springer-Verlag, New York.
- De Cesare, L., Myers, D. E., Posa, D., 2001. Estimating and modelling space-time correlation structures. *Stat. Probab. Lett.* 51, 9–14.
- De Iaco, S., Myers, D. E., Posa, D., 2001. Space-time analysis using a general product-sum model. *Stat. Probab. Lett.* 52, 21–28.
- De Iaco, S., Myers, D. E., Posa, D., 2002. Space-time variograms and a functional form for total air pollution measurements. *Comput. Stat. Data Anal.* 41, 311–328.
- Desmond, A. F., 1997. Optimal estimating functions, quasi-likelihood and statistical modelling. *J. Stat. Plan. Inference* 60, 77–121.
- ESYRCE (1995–2002). Encuesta de superficies y rendimientos de cultivos, 1995–2002. Subdirección General de Estadísticas Agroalimentarias. Secretaría General Técnica. Ministerio de Agricultura, Pesca y Alimentación, Madrid.
- Fahrmeir, L., Kaufmann, H., 1987. Regression models for non-stationary categorical time series. *J. Time Ser. Anal.* 8, 147–160.
- Fingleton, B., 1988. Categorical data with inherent spatial dependence: the case of cluster sampling. *Trans. Inst. Br. Geogr.* 13, 497–503.
- Gallego, F. J., Delincé, J., Carfagna, E., 1994. Two stage area frame sampling on square segments for farm surveys. *Surv. Methodol.* 20, 107–116.
- Godambe, V. P., Thompson, M. E., 1989. An extension of quasi-likelihood estimation. *J. Stat. Plan. Inference* 22, 137–172.
- Golan, A., Judge, G., Miller, D., 1996. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. John Wiley, New York.
- Gotway, C. A., Young, L. J., 2002. Combining incompatible spatial data. *J. Am. Stat. Assoc.* 97, 633–648.
- Gourieroux, C., Monfort, A., Trognon, A., 1984. Pseudo maximum likelihood methods: applications to Poisson models. *Econometrica* 52, 701–720.

- Grömping, U., 1996. A note on fitting a marginal model to mixed effects log-linear regression data via GEE. *Biometrics* 52, 280–285.
- Hardie, I. W., Parks, P. J., 1997. Land use with heterogeneous land quality: an application of an area base model. *Am. J. Agric. Econ.* 79, 299–310.
- Hart, P. W. E., 1980. Problems and potentialities of behavioural approach to agricultural location. *Geogr. Ann. Ser. B.* 62, 99–107.
- Howit, R., Reynaud, A., 2003. Spatial disaggregation of agricultural production data using maximum entropy. *Eur. Rev. Agric. Econ.* 30, 359–387.
- Hsiao, C., 2003. *Analysis of Panel Data*. Cambridge University Press; Cambridge, UK.
- Ijiri, Y., 1971. Fundamental queries in aggregation theory. *J. Am. Stat. Assoc.* 66, 766–782.
- Irwin, F., Geoghegan, J., 2001. Theory, data, and methods: developing spatially explicit economic models of land use change. *Agric. Ecosys. Environ.* 85, 7–24.
- Lang, J. B., 1996. On the comparison of multinomial and Poisson log-linear models. *J. Royal Stat. Soc. Ser. B.* 58, 253–266.
- Lee, A. H., Zhao, Y., 1997. Assessing influence on the goodness-of-link in generalized linear models. *Stat. Probab. Lett.* 31, 351–358.
- Liang, K. Y., Zeger, S., 1986. Longitudinal data analysis using generalized linear models. *Biometrika* 73, 13–22.
- Lichtenberg, E., 1989. Land quality, irrigation development and cropping patterns in the Northern High Plains. *Am. J. Agric. Econ.* 83, 705–710.
- Luijten, J. C., 2003. A systematic method for generating land use patterns using stochastic rules and basic landscape characteristics: results for a Colombian hillside watershed. *Agric. Ecosys. Environ.* 95, 427–441.
- Luke, D. A., 2004. *Multilevel Modelling. Series: Quantitative applications in the Social Sciences*. Sage University Papers, Thousand Oaks, CA.
- McCullagh, P., 1983. Quasi-likelihood functions. *Ann. Stat.* 11, 59–67.
- McCulloch, C. E., Searle, S. R., 2001. *Generalized, Linear, and Mixed Models*. Wiley, New York.
- McShane, L. M., Albert, P. S., Palmatier, M. A., 1997. A latent process regression model for spatially correlated count data. *Biometrics* 53, 698–706.
- Miller, D. J., Plantinga, A. J., 1999. Modelling land use decisions with aggregate data. *Am. J. Agric. Econ.* 81, 180–194.
- Müller, D., Zuller, M., 2002. Land use dynamics in the central highlands of Vietnam: A spatial model combining village survey data with satellite imagery interpretation. *Agric. Econ.* 27, 333–354.
- Munroe, D. K., Southworth, J., Tucker, C. M., 2002. The dynamics of land-cover change in western Honduras: exploring spatial and temporal complexity. *Agric. Econ.* 27, 355–369.
- Nelson, G. C., 2002. Introduction to the special issue on spatial analysis for agricultural economics. *Agric. Econ.* 27, 197–200.
- Parks, P. J., Murray, B. C., 1994. Land attributes and land allocation: non-industrial forest use in the Pacific Northwest. *For. Sci.* 40, 558–575.
- Pickering, N. B., Jones, J. W., Boote, K. J., 1995. Adapting SOYGRO V5.42 for predicting under climate change conditions. In: Rosenzweig, C., Ritchie, J. T., Jones, J. W., Tsuji, G. Y., Hildebrand, P. J. (Eds.) *Climate Change and Agriculture: Analysis of Potential International Impacts*. ASA Spec. Publ. 59. ASA, CSSA, and SSSA, Madison, WI.
- Plantinga, A. J., 1996. The effect of agricultural policies on land use and environmental quality. *Am. J. Agric. Econ.* 78(4), 1082–1091.
- Plantinga, A. J., Buonjorno, J., Alig, R. J., 1990. Determinant of changes in non-industrial private timberland ownership in the United States. *J. World For. Res. Manage.* 5, 29–46.
- Polsky, C., Earsterling, W. E., 2001. Adaptation to climate variability and change in the US Great Plains: a multi-scale analysis of Ricardian climate sensitivities. *Agric. Ecosys. Environ.* 85, 133–144.
- Prentice, R. L., 1988. Correlated binary regression with covariates specific to each binary observation. *Biometrics* 44, 1033–1048.
- Robinson, G. K., 1991. That BLUP is a good thing: the estimation of random effects. *Stat. Sci.* 6, 15–51.
- SAS, 1997. SAS Institute Inc., SAS/STAT Software: Changes and Enhancements through Release 6.12, SAS Institute Inc., Cary, NC, U.S.A.
- Skrondal, A., Rabe-Hesketh, S., 2004. *Generalized Latent Variable Modelling*. Chapman&Hall/CRC, Boca Raton, FL.
- Stavins, R. N., Jaffe, A. B., 1990. Unintended impacts of public investments on private decisions: the depletion of forest wetlands. *Am. Econ. Rev.* 80, 337–352.
- Stoker, T. M., 1993. Empirical approaches to the problem of aggregation over individuals. *J. Econ. Lit.* 31, 1827–1874.
- Valliant, R., 1985. Non-linear prediction theory and the estimation of a proportion in a finite population. *J. Am. Stat. Assoc.* 80, 631–641.
- Valliant, R., 1986. Mean squared error estimation in finite population under non-linear models. *Commun. Stat. A* 15, 1975–1993.
- van Raaij, B. F. T., van der Wal, T., 1994. CGMS version 3.1. Technical description. JRC/SC-DLO/Quay. Technical Document 15.3. Wageningen, The Netherlands.
- Veen, A., Otter, H. S., 2001. Land uses changes in regional economic theory. *Environ. Model. Assess.* 6, 145–150.
- Vijapurkar, U. P., Gotway, C. A., 2001. Assessment of forecast and forecast uncertainty using generalized linear regression models for time series count data. *J. Stat. Comput. Simul.* 68, 321–349.
- Wedderburn, R. W. M., 1974. Quasi-likelihood functions, generalized linear models and the Gauss–Newton method. *Biometrika* 61, 439–447.
- Whisler, F. D., Acock, B., Baker, D. N., Fye, R. E., Hodges, H. F., Lambert, J. R., Lemman, H. E., McKinion, J. M., Reddy, V. R., 1986. Crop simulations models in agronomic systems. *Adv. Agron.* 40, 141–208.
- White, F. C., Fleming, F. N., 1980. An analysis of competing agricultural land uses. *South. J. Agric. Econ.* 12, 99–103.
- Wu, J., Segerson, K., 1995. The impact of policies and land characteristics on potential groundwater pollution in Wisconsin. *Am. J. Agric. Econ.* 77(4), 1033–1047.
- Wu, J., Adams, R. M., Kling, C. L., Tanaka, K., 2004. From micro-level decisions to landscape changes: an assessment of agricultural conservation policies. *Am. J. Agric. Econ.* 88(1), 26–41.
- Zeger, S. L., 1988. A regression model for time series of counts. *Biometrika* 75, 621–629.
- Zeger, S. L., Liang, K. Y., 1986. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 42, 121–130.
- Zeger, S. L., Liang, K. Y., Albert, P. S., 1988. Models for longitudinal data: a generalized estimating equations approach. *Biometrics* 44, 1049–1060.
- Zimmerman, D. L., Zimmerman, M. B., 1991. A comparison of spatial semi-variogram estimators and corresponding ordinary Kriging predictors. *Technometrics* 33, 77–91.